

DOCUMENT RESUME

ED 275 702

TM 860 609

AUTHOR Samejima, Fumiko
TITLE Effect of the Guessing Parameter on the Estimation of the Item Discrimination and Difficulty Parameters When Three-Parameter Logistic Model Is Assumed.
PUB DATE Apr 86
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Simulation; Difficulty Level; *Guessing (Tests); *Item Analysis; *Latent Trait Theory; Mathematical Models; Monte Carlo Methods; Multiple Choice Tests; *Statistical Bias; Statistical Studies; Test Items; Test Length
IDENTIFIERS Logist Computer Program; *Normal Ogive Models; *Rasch Model

ABSTRACT

Item analysis data fitting the normal ogive model were simulated in order to investigate the problems encountered when applying the three-parameter logistic model. Binary item tests containing 10 and 35 items were created, and Monte Carlo methods simulated the responses of 2,000 and 500 examinees. Item parameters were obtained using Logist 5. Results indicated a substantial effect of the assumed third parameter on the other two estimated item parameters, if the estimation was made by molding the item characteristic function into the three parameter model when it actually followed the normal ogive model. This effect was stronger on the estimated item discrimination parameter than the difficulty parameter. Both of these parameters tended to be overestimated in order to counteract the effect of the guessing parameter. To amend these enhancements, the discrimination shrinkage factor and the difficulty reduction index were proposed. Results also indicated that tests must contain about 80 items in order to make good use of Logist 5. There was no substantial difference between 500 or 2,000 subjects. It was concluded that it may be dangerous to assume the validity of the three-parameter model, especially when test items are relatively easy. (GDC)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

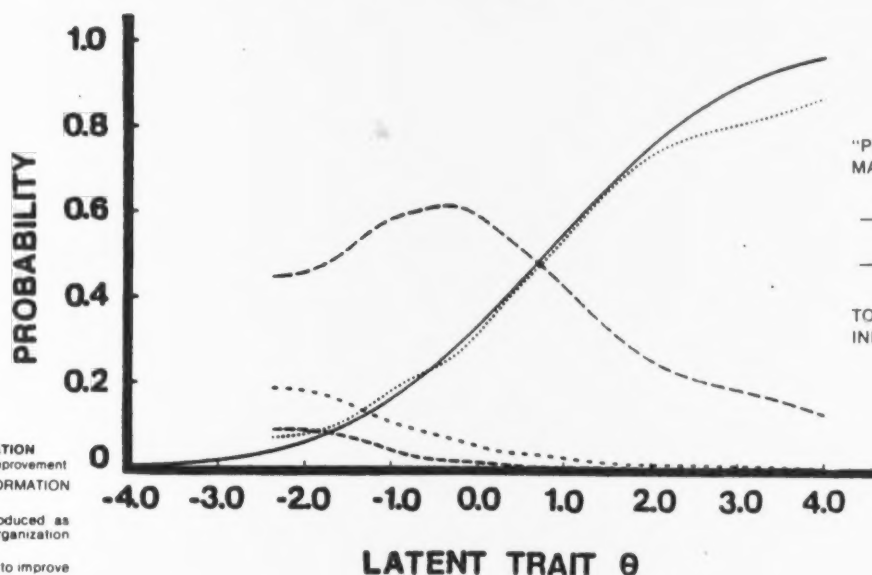
ED275702

1986 AERA Annual Meeting
San Francisco

Fumiko Samejima
University of Tennessee

Effect of the Guessing Parameter on the Estimation of the Item
Discrimination and Difficulty Parameters When Three-Parameter Logistic
Model Is Assumed

Three-parameter logistic model has been used for the multiple-choice test item by many researchers. Certain non-parametric estimation of the operating characteristics of the wrong alternative answers has revealed, however, many multiple-choice test items do not follow the model, nor do they follow the Equivalent Distractor Model in general, to which the three-parameter logistic model belongs. Those items can best be interpreted by the Informative Distractor Model. Figure 1 shows an example of the set of operating characteristics of the four alternative answers to an item taken from the Level 12 Vocabulary Test of the Iowa Test of Basic Skills, which were estimated by the Simple Sum Procedure of the Conditional P.D.F.



"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

F. Samejima

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✖ This document has been reproduced as
received from the person or organization
originating it.

□ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

FIGURE 1

Approach combined by the Normal Approach Method (Samejima, 1981). We can see in this graph that each distractor has its own unique operating

FM860609

characteristic, or plausibility function, and also the estimated item characteristic function is fairly close to the one in the normal ogive model, which is drawn by a solid line in the figure. This set of operating characteristics can better be represented by one of the family of models proposed for the multiple-choice test item, which takes account of the unique information provided by each distractor as well as the effect of the examinee's random guessing behavior (cf. Samejima, 1979). Figure 2 illustrates the item characteristic function in Model A. We can see that

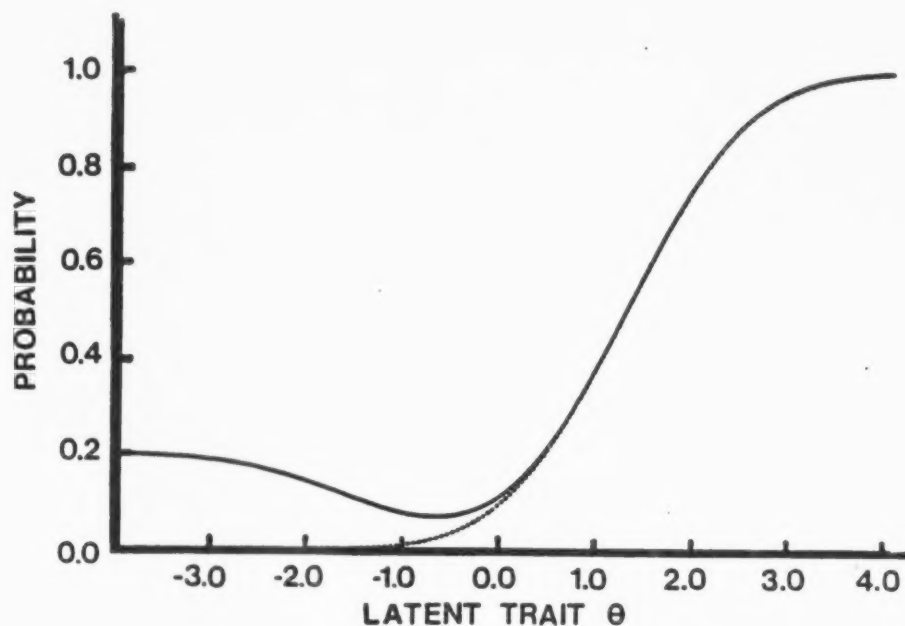


FIGURE 2

it is very close to the one in the normal ogive model which is drawn by a dotted line, except for the lower part of the curve, the conditional probability of success which is almost entirely caused by random guessing. In cases like this, it will be wise to approximate the curve by the normal

ogive function by discarding the item response in estimating lower ability, since it provides us with nothing but noise.

For the reasons described above, the normal ogive model has its own place in the multiple-choice test item, since it often fits the item better than the three-parameter logistic model. It will be worthwhile, therefore, to investigate and clarify the problems encountered when we apply the three-parameter logistic model to the data which actually follow the normal ogive model. The data used in the present study are simulated data. There are basically two sets of hypothetical test items, all of which follow the normal ogive model. The first set consists of ten binary test items, whose item discrimination and difficulty parameters, a_g and b_g , for each of these ten items are shown in Tables 1 and 3. The second test consists of thirty-five binary test items, and their discrimination and difficulty

TABLE 1

Theoretical and Estimated Item Parameters in the 500 and 2,000 Subject Cases for Each Item of the Ten Item Test. Three-Parameter Logistic Model Is Assumed. Case 1.

Item	Discrimination Parameter				Difficulty Parameter				Guessing Parameter	
	Theoretical		Estimated		Theoretical		Estimated		Estimated	
	Org.	Adj.	500 S.C.	2,000 S.C.	Org.	Adj.	500 S.C.	2,000 S.C.	500 S.C.	2,000 S.C.
1	1.50000	2.16506	4.00000	0.95615	-2.50000	-1.73205	-7.00850	-4.53097	0.11111	0.00000
2	1.00000	1.44338	1.12847	1.89857	-2.00000	-1.38564	-2.27705	-1.25111	0.11111	0.23137
3	2.50000	3.60844	4.00000	7.00000	-1.50000	-1.03923	-1.49140	-1.02797	0.01285	0.01897
4	1.00000	1.44338	1.49368	2.16571	-1.00000	-0.69282	-0.76647	-0.34862	0.12154	0.15474
5	1.50000	2.16506	4.00000	5.73384	-0.50000	-0.34641	-0.39155	-0.07764	0.01397	0.04219
6	1.00000	1.44338	1.81350	2.45766	0.00000	0.00000	-0.00846	0.25474	0.04756	0.06358
7	2.00000	2.88675	3.08340	6.30172	0.50000	0.34641	0.32296	0.49745	0.00000	0.00499
8	1.00000	1.44338	1.47734	2.11129	1.00000	0.69282	0.72304	0.81349	0.00000	0.00645
9	2.00000	2.88675	4.00000	5.73230	1.50000	1.03923	1.10169	1.04610	0.00000	0.00000
10	1.00000	1.44338	0.65759	0.98121	2.00000	1.38564	2.60342	2.08713	0.00000	0.00000

TABLE 2

Theoretical and Estimated Item Parameters in the 500 and 2,000 Subject Cases for Each Item of the Thirty-Five Item Test. Three-Parameter Logistic Model Is Assumed. Case 2.

Item	Discrimination Parameter				Difficulty Parameter				Guessing Parameter	
	Theoretical		Estimated		Theoretical		Estimated		Estimated	
	Org.	Adj.	500 S.C.	2,000 S.C.	Org.	Adj.	500 S.C.	2,000 S.C.	500 S.C.	2,000 S.C.
11	1.80000	2.59808	---	---	-4.75000	-3.29090	---	---	---	---
12	1.90000	2.74241	---	---	-4.50000	-3.11769	---	---	---	---
13	2.00000	2.88675	---	---	-4.25000	-2.94449	---	---	---	---
14	1.50000	2.16506	7.00000	2.10115	-4.00000	-2.77128	-2.31556	-3.02161	0.08667	0.02122
15	1.60000	2.30940	---	4.95191	-3.75000	-2.59808	---	-2.41505	---	0.02122
16	1.40000	2.02073	7.00000	2.01814	-3.50000	-2.42487	-2.31184	-2.66637	0.08667	0.02122
17	1.90000	2.74241	2.34600	6.25083	-3.00000	-2.07846	-2.49184	-1.97065	0.08667	0.00000
18	1.80000	2.59808	1.66937	2.45684	-3.00000	-2.07846	-2.53506	-2.23919	0.08667	0.02122
19	1.60000	2.30940	5.95996	2.62153	-2.75000	-1.90526	-1.63507	-1.90260	0.33907	0.02122
20	2.00000	2.88675	4.54064	3.93175	-2.50000	-1.73205	-1.75362	-1.66113	0.00000	0.16847
21	1.50000	2.16506	3.60897	3.57269	-2.25000	-1.55885	-1.52887	-1.47593	0.00000	0.00000
22	1.70000	2.45374	3.72440	5.79677	-2.00000	-1.38564	-1.26918	-1.20411	0.19618	0.17118
23	1.50000	2.16506	5.93530	4.14441	-1.75000	-1.21244	-0.86540	-1.04556	0.31164	0.12398
24	1.40000	2.02073	2.15613	2.50797	-1.50000	-1.03923	-0.91905	-0.92924	0.05498	0.08893
25	2.00000	2.88675	4.30412	3.82653	-1.25000	-0.86603	-0.82709	-0.84287	0.00000	0.01227
26	1.60000	2.30940	3.08081	3.00092	-1.00000	-0.69282	-0.61492	-0.62173	0.00000	0.03369
27	1.80000	2.59808	3.41671	3.29292	-0.75000	-0.51962	-0.42482	-0.50742	0.01715	0.00712
28	1.70000	2.45374	3.73975	3.15448	-0.50000	-0.34641	-0.29610	-0.32412	0.00000	0.00000
29	1.90000	2.74241	3.55043	3.43681	-0.25000	-0.17321	-0.16918	-0.15937	0.00000	0.00000
30	1.70000	2.45374	3.26689	2.97724	0.00000	0.00000	0.06045	0.03331	0.00000	0.00000
31	1.50000	2.16506	2.51771	2.68637	0.25000	0.17321	0.24396	0.17012	0.00000	0.00132
32	1.80000	2.59808	3.36240	3.58997	0.50000	0.34641	0.38959	0.35692	0.00000	0.00000
33	1.40000	2.02073	2.49348	2.33960	0.75000	0.51962	0.49717	0.47978	0.00000	0.00000
34	1.90000	2.74241	3.92561	3.63878	1.00000	0.69282	0.64148	0.65455	0.00000	0.00000
35	2.00000	2.88675	4.36628	3.48851	1.25000	0.86603	0.78414	0.85172	0.00000	0.00000
36	1.60000	2.30940	2.61472	2.72612	1.50000	1.03923	1.01803	1.02596	0.00000	0.00000
37	1.70000	2.45374	2.54991	2.90554	1.75000	1.21244	1.21299	1.22349	0.00000	0.00000
38	1.40000	2.02073	1.97089	2.25224	2.00000	1.38564	1.38765	1.34179	0.00000	0.00000
39	1.90000	2.74241	6.12277	3.14405	2.25000	1.55885	1.47525	1.52861	0.07920	0.00000
40	1.60000	2.30940	4.30567	3.39256	2.50000	1.73205	1.51061	1.62665	0.00000	0.00000
41	1.50000	2.16506	2.50260	2.18581	2.75000	1.90526	2.09183	1.97325	0.00000	0.00000
42	1.70000	2.45374	7.00000	6.00998	3.00000	2.07846	1.82767	1.88083	0.00000	0.00000
43	1.80000	2.59808	3.18005	1.91086	3.25000	2.25167	2.10989	2.68272	0.00000	0.00000
44	2.00000	2.88675	---	---	3.50000	2.42487	---	---	---	---
45	1.40000	2.02073	---	7.00000	3.75000	2.59808	---	2.19254	---	0.00198

TABLE 3

Theoretical and Estimated Item Parameters in the 500 and 2,000 Subject Cases for Each Item of the Ten Item Test and the Thirty-Five Item Test. Three-Parameter Logistic Model Is Assumed. Case 3.

Item	Discrimination Parameter				Difficulty Parameter				Guessing Parameter	
	Theoretical		Estimated		Theoretical		Estimated		Estimated	
	Org.	Adj.	500 S.C.	2,000 S.C.	Org.	Adj.	500 S.C.	2,000 S.C.	500 S.C.	2,000 S.C.
1	1.50000	2.16506	7.00000	3.16857	-2.50000	-1.73205	-1.58337	-1.70186	0.31911	0.00000
2	1.00000	1.44338	3.87013	2.20593	-2.00000	-1.38564	-0.82818	-0.99213	0.42360	0.33765
3	2.50000	3.60844	6.61387	6.11440	-1.50000	-1.03923	-0.88235	-0.94942	0.07876	0.06904
4	1.00000	1.44338	1.66601	1.60446	-1.00000	-0.69282	-0.57784	-0.65341	0.06874	0.02886
5	1.50000	2.16506	2.77616	2.49209	-0.50000	-0.34641	-0.27650	-0.33620	0.00000	0.00000
6	1.00000	1.44338	1.55749	1.58121	0.00000	0.00000	0.02360	-0.00453	0.01465	0.00128
7	2.00000	2.88675	2.96359	3.40269	0.50000	0.34641	0.36510	0.36758	0.00000	0.00022
8	1.00000	1.44338	1.76403	1.68448	1.00000	0.69282	0.65648	0.69709	0.00000	0.00000
9	2.00000	2.88675	4.38452	3.69650	1.50000	1.03923	0.97398	0.98698	0.00000	0.00000
10	1.00000	1.44338	1.54466	1.67432	2.00000	1.38564	1.32856	1.35974	0.00000	0.00000
11	1.80000	2.59808	---	---	-4.75000	-3.29090	---	---	---	---
12	1.90000	2.74241	---	---	-4.50000	-3.11769	---	---	---	---
13	2.00000	2.88675	---	---	-4.25000	-2.94449	---	---	---	---
14	1.50000	2.16506	7.00000	2.19806	-4.00000	-2.77128	-2.37410	-2.95371	0.08826	0.10234
15	1.60000	2.30940	---	1.90897	-3.75000	-2.59808	---	-3.36769	---	0.10234
16	1.40000	2.02073	7.00000	1.84048	-3.50000	-2.42487	-2.37226	-2.75675	0.08826	0.10234
17	1.90000	2.74241	2.00167	5.60182	-3.00000	-2.07846	-2.63823	-1.99530	0.08826	0.00000
18	1.80000	2.59808	1.62591	2.21154	-3.00000	-2.07846	-2.57808	-2.29576	0.08826	0.10234
19	1.60000	2.30940	4.45727	2.36176	-2.75000	-1.90526	-1.63765	-1.92579	0.43106	0.10234
20	2.00000	2.88675	6.38769	3.87045	-2.50000	-1.73205	-1.73779	-1.68322	0.00000	0.14877
21	1.50000	2.16506	2.89572	3.06226	-2.25000	-1.55885	-1.57782	-1.50616	0.00000	0.00000
22	1.70000	2.45374	3.03416	4.41425	-2.00000	-1.38564	-1.33087	-1.18774	0.14331	0.20630
23	1.50000	2.16506	3.81430	3.36983	-1.75000	-1.21244	-0.89510	-1.05681	0.28760	0.10975
24	1.40000	2.02073	2.17074	2.43977	-1.50000	-1.03923	-0.87115	-0.92907	0.09159	0.08490
25	2.00000	2.88675	3.70917	3.62242	-1.25000	-0.86603	-0.81486	-0.83289	0.00000	0.01472
26	1.60000	2.30940	2.86861	2.96961	-1.00000	-0.69282	-0.59335	-0.60991	0.00000	0.03712
27	1.80000	2.59808	3.06301	3.14587	-0.75000	-0.51962	-0.41924	-0.50798	0.00089	0.00000
28	1.70000	2.45374	4.17905	3.18564	-0.50000	-0.34641	-0.27372	-0.31965	0.00000	0.00000
29	1.90000	2.74241	3.65756	3.30394	-0.25000	-0.17321	-0.15748	-0.15832	0.00000	0.00000
30	1.70000	2.45374	3.36662	2.92438	0.00000	0.00000	0.05765	0.03296	0.00000	0.00000
31	1.50000	2.16506	2.49791	2.60306	0.25000	0.17321	0.24144	0.16800	0.00000	0.00000
32	1.80000	2.59808	2.97532	3.17624	0.50000	0.34641	0.38466	0.35806	0.00000	0.00000
33	1.40000	2.02073	2.39944	2.24392	0.75000	0.51962	0.49468	0.48257	0.00000	0.00000
34	1.90000	2.74241	3.56410	3.65448	1.00000	0.69282	0.64140	0.65986	0.00000	0.00000
35	2.00000	2.88675	4.31324	3.29432	1.25000	0.86603	0.78225	0.85445	0.00000	0.00000
36	1.60000	2.30940	2.51338	2.56816	1.50000	1.03923	1.01267	1.02876	0.00000	0.00000
37	1.70000	2.45374	2.43304	2.70841	1.75000	1.21244	1.20847	1.22648	0.00000	0.00000
38	1.40000	2.02073	1.99664	2.25976	2.00000	1.38564	1.37635	1.33734	0.00000	0.00000
39	1.90000	2.74241	2.91426	3.11288	2.25000	1.55885	1.47260	1.52379	0.00000	0.00000
40	1.60000	2.30940	3.98169	3.20664	2.50000	1.73205	1.51130	1.63052	0.00000	0.00000
41	1.50000	2.16506	2.40751	2.03311	2.75000	1.90526	2.12160	2.00855	0.00000	0.00000
42	1.70000	2.45374	7.00000	6.06362	3.00000	2.07846	1.86710	1.88062	0.00000	0.00000
43	1.80000	2.59808	2.68843	1.73795	3.25000	2.25167	2.20918	2.79685	0.00000	0.00000
44	2.00000	2.88675	---	---	3.50000	2.42487	---	---	---	---
45	1.40000	2.02073	---	7.00000	3.75000	2.59808	---	2.19260	---	0.00198

parameters are shown in Table 2 and 3. Thus we have forty five binary test items in total. For brevity, hereafter, we shall call these two sets of hypothetical binary test items Ten Item Test and Thirty-Five Item Test, respectively.

Five hundred examinees whose ability distributes uniformly within the interval of θ , $(-2.5, 2.5)$, were hypothesized. Actually, there are one hundred discrete points of θ ranging from -2.475 to 2.475 with the equal step of 0.050 , and five hypothetical subjects are placed at each ability level. The response pattern for each of these five hundred hypothetical subjects was produced by Monte Carlo Method. In order to investigate the effect of the number of test items on the resultant estimated parameters obtained by Logist 5, we used: 1) the first ten items, 2) the second thirty-five items, and 3) the forty-five items combining these two sets of items, separately.

As an additional observation, we wished to increase the number of items in the total set of binary items for which Logist 5 was to be applied. Thus another set of thirty-five binary test items was added more or less artificially (cf. Samejima, 1984). Thus in addition to the three cases described earlier, we have: 4) the eighty items adding those "artificial" thirty-five items to the forty-five items. Hereafter, we shall call these four different situations Cases 1, 2, 3, and 4, respectively.

For the purpose of investigating the effect of the number of examinees on the resultant estimated parameters obtained by Logist 5, an additional group of one thousand five hundred examinees were further hypothesized, to provide us with a larger group of two thousand examinees, whose ability follows the same uniform distribution. Thus in this larger

group twenty examinees are placed at each of the one hundred equally spaced discrete ability levels, which were described earlier. The response pattern of the forty-five binary test item scores and the additional thirty-five "artificial" binary test item scores was produced for each of the one thousand five hundred subjects, as we did for each of the five hundred subjects. For brevity, hereafter we shall call these two situations 500 Subject Case and 2,000 Subject Case, respectively, depending upon the number of hypothetical examinees included in the sample.

In the results of Logist 5, the origin and the unit of the ability scale are adjusted to the mean and the standard deviation of the sample distribution of the estimated individual ability parameters of the examinees, which did not exceed 3.0 in absolute values after the last iteration. The theoretical item and individual parameters were transformed in order to make them comparable to the estimated parameters obtained by Logist 5. Those adjusted theoretical item parameters are also shown in Tables 1, 2, and 3. On the new scale of θ the true individual ability parameters of our hypothetical examinees are placed at one hundred equally spaced ability levels starting from approximately -1.714730299 and ending with 1.714730299, with the equal step of approximately 0.034641016.

The resulting estimated item parameters in Case 1 obtained by Logist 5, in both 500 and 2,000 Subject Cases, are presented in Table 1 for each item, and those for Cases 2 and 3 are in Tables 2 and 3, respectively.

These results suggest that there exists a substantial effect of the assumed third parameter, c_g , on the other two estimated item parameters, if the estimation is made by "molding" the item characteristic function into that of the three-parameter logistic model, when actually it follows the normal ogive model. This effect appears to be stronger on the estimated discrimination parameter than on the estimated difficulty

parameter.

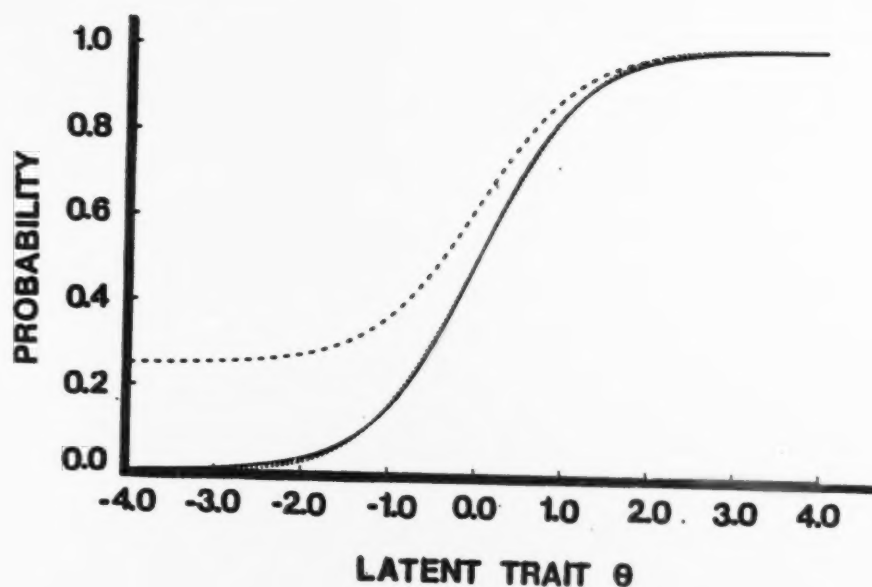


FIGURE 3

Figure 3 presents the item characteristic function in the normal ogive model with $a_g = 1.00$ and $b_g = 0.00$ by a dotted line, the one in the logistic model with the same parameters and the scaling factor, $D = 1.7$, by a solid line, and the one in the three-parameter logistic model with the same two item parameters and scaling factor and the third parameter, $c_g = 0.25$, by a dashed line. It is obvious from theory that for all the three item characteristic functions the derivative is highest at

$\theta = b_g = 0.00$. Actually, these three derivatives are: $(2\pi)^{-1/2}a_g$, $Da_g/4$ and $(1-c_g)Da_g/4$, respectively, for the three functions in Figure 3. The ratio of this maximal slope in the normal ogive model to the one in the logistic model is approximately 0.938687718, which is not so much less than unity. The corresponding ratio between the three-parameter logistic model and the logistic model is $(1-c_g)$, which equals 0.75 when

$c_g = 0.25$, and is as low as 0.50 when $c_g = 0.50$. The ratio between the three-parameter logistic model and the normal ogive model is approximately $1.065317017(1-c_g)$, which is a little higher than $(1-c_g)$.

It is obvious from the facts described above that, when we "mold" the normal ogive model into the three-parameter logistic model and estimate the three item parameters accordingly, we tend to obtain a larger value for the estimated item discrimination parameter, in order to "fight" the effect of the guessing parameter. It is especially so when the estimated guessing parameter turns out to be a large value for one reason or another. It is also obvious from Figure 3 that the difficulty parameter tends to be estimated higher than it really is. In order to amend these enhancements, the discrimination shrinkage factor and the difficulty reduction index are proposed by formulae (2) and (4), respectively.

$$(1) \quad a_g^* = \zeta(c_g^*) a_g \quad .$$

$$(2) \quad \zeta(c_g^*) = \{-\log(1-2c_g^*)\} \{\log(1+c_g^*) - \log(1-c_g^*)\}^{-1} \quad .$$

$$(3) \quad b_g^* = b_g + \xi(c_g^*|a_g) \quad .$$

$$(4) \quad \xi(c_g^*|a_g) = (Da_g)^{-1} \{\log(1+c_g^*) - \log(1-c_g^*)\} \quad .$$

In these formulae, a_g^* , b_g^* , and c_g^* indicate the estimated item discrimination, difficulty and guessing parameters when the three-parameter logistic model is assumed, respectively.

Some resulting estimated item characteristic functions are disastrously different from the theoretical item characteristic functions,

especially in Case 1 where only ten binary test items were included. This seems largely due to the fact that in Logist 5, in estimating the item parameters, the estimated individual parameter, which is based upon the maximum likelihood estimation, is treated as if it were the true individual parameter, and its large error of ability estimation caused by the small number of items affects the item parameter estimation, especially near the endpoints of the interval of θ for which our subjects' ability distributes. Thus the estimated item characteristic functions of very easy test items become disastrously different from the true item characteristic functions. Since there is no functional relationship between θ and its maximum likelihood estimate $\hat{\theta}$, there is no simple way to correct these biases. We must keep in mind that in using Logist 5, or any other program based upon the same principle, the number of test items must be substantially large, and the items whose parameters are to be estimated must be of intermediate difficulty relative to the ability distribution. It appears that forty-five is not a large enough number of items to be included, and that we need at least some number in the vicinity of eighty, in order to make a good use of Logist 5. Otherwise, we will obtain false "tails" caused by large values of the estimated guessing parameter, which are supposed to be zero. The problem of the enhancement of the estimated item discrimination parameters cannot completely be solved by using the discrimination shrinkage factor. Since in the results of Logist 5 the mean and the standard deviation of the maximum likelihood estimate of ability, instead of those of ability themselves, are used as the origin and the unit of θ , respectively, those inflated estimates of a_g indicate that the standard deviation of the maximum likelihood estimate may actually be larger than that of ability, in each of the sixteen different situations. This will be observed further and discussed later in

... σ_g indicate that the standard deviation of the maximum likelihood estimate may actually be larger than that of ability, in each of the sixteen different situations. This will be observed further and discussed later in a separate paper.

We find no substantial differences between the results of 500 Subject Case and 2,000 Subject Case. This indicates that increasing the number of subjects from 500 to 2,000 does not provide us with a substantial gain.

The results of the present research imply the warning that it is dangerous to assume three-parameter logistic model when we have no evidence for its validity, especially when the items are relatively easy in comparison with the ability distribution. It may provide us with large "tails" in the estimated item characteristic functions, when actually they do not exist.

References

- [1] Samejima, F. A new family of models for the multiple-choice item. Office of Naval Research Report 79-4, 1979.
- [2] Samejima, F. Final Report: Efficient methods of estimating the operating characteristic of item response categories and challenge to a new model for the multiple-choice item. Final Report of N0014-77-C-0360, Office of Naval Research, 1981.
- [3] Samejima, F. Results of item parameter estimation using logist 5 on simulated data. Office of Naval Research Report 84-3, 1984.